

ПРИМЕНЕНИЕ СОРТОВОЙ СИСТЕМЫ ЕСТЕСТВЕННОГО ЯЗЫКА В ЗАДАЧЕ УСТАНОВЛЕНИЯ СМЫСЛОВОЙ ЭКВИВАЛЕНТНОСТИ ТЕКСТОВ.

Д.В.Михайлов, Г.М.Емельянов

Цель. исследование практических аспектов построения теории Лексического Значения слова по его словарному определению для автоматизации выделения свойств обозначаемого словом объекта и формирования в полуавтоматическом режиме множества аксиом теории.

Задачи исследования.

- 1) Разработать реализуемое программно формальное описание теории Лексического Значения слова с ориентацией на общие принципы использования семантической информации при построении дерева синтаксического подчинения и полученную ранее структуру языковой базы знаний для задачи установления семантической эквивалентности текстов Естественного Языка.
- 2) На основе введенного формального представления разработать методику систематизации и контроля корректности наборов утверждений-аксиом, используемых при построении теорий.



Булкин Паша - Тест №1 'Физика' (2 вопр)

[Отмена](#)[Сдать на проверку](#)

Вопрос 1	Вопрос 1
Ответ	
Вопрос 2	Что происходит согласно 1 закону Ньютона с материальным телом, на которое не действуют внешние силы ?
Ответ	

Ввод ответа

Что происходит согласно 1 закону Ньютона с материальным телом, на которое не действуют внешние силы ?

Тело остается в покое или совершает движение с постоянной скоростью по прямой.

OK

Отмена

Тест открытой формы : соотнесение смыслов.

Вопрос 1 : Что происходит согласно 1-му закону Ньютона с материальным телом, на которое не действуют внешние силы ?

Эталонный ответ для вопроса 1 : Тело покоится или движется прямолинейно и равномерно.

Введенный правильный ответ : Тело сохраняет покой или движется прямолинейно и равномерно.

Смысловое варьирование :

$Покоиться \Leftrightarrow Покой \xleftarrow{2} Сохранять$

что эквивалентно записи с применением обозначений Лексических Функций :

$Покоиться \Leftrightarrow S_0(Покоиться) \xleftarrow{2} Oper_1(S_0(Покоиться))$

Подобная замена соответствует лексическому правилу :

$C_0 \Leftrightarrow S_0(C_0) \xleftarrow{2} Oper_1(S_0(C_0))$

(правило № 17 из описанных И.А. Мельчуком)

Вопрос 2 : Каким образом взаимодействуют друг с другом провода контактной сети троллейбусной линии ?

Эталонный ответ для вопроса 2 : Отталкиваются в силу противоположности направления движения тока.

Введенный правильный ответ : Отталкиваются как следствие противоположности направления движения тока.

Смысловое варьирование :

$в силу \Leftrightarrow как следствие$

что эквивалентно записи с применением обозначений Лексических Функций :

$в силу \Leftrightarrow Syn(в силу)$

(правило № 1 из описанных И.А. Мельчуком)

Неоднозначность выбора ситуации при построении ГСС анализируемой фразы.

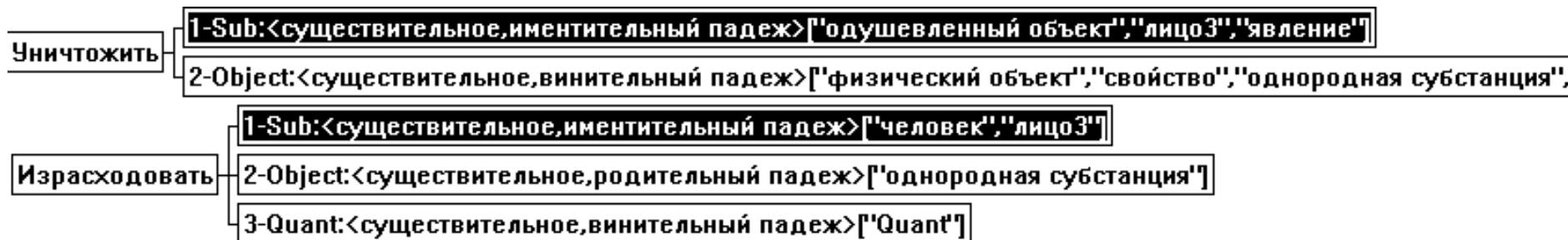


Рис.1. Модель Управления для глагола *сжечь* в значениях *уничтожить* и *израсходовать*.

Возможная неадекватная перифраза :

За зиму он сжест $[C_0]$ *машину дров* \Leftrightarrow
 \Leftrightarrow *За зиму он уничтожил* $[Oper_1(S_{instr}(C_0))]$
огнем $[S_{instr}(C_0)]$ *машину дров.*

(правило № 24 из описанных И.А. Мельчуком)

Пример использования синтагматического ограничения на применимость правила при разрешении неоднозначности выбора ситуации.

За зиму он сжег $\xrightarrow{5}$ *по* $[Pr opt(неосторожность)] \xrightarrow{2}$ *неосторожности* $[неосторожность]$
машину дров.

Именная Группа *по неосторожности* $[Pr opt(неосторожность)] \xrightarrow{2}$ *неосторожность* при сказуемом *сжег* выражает адьюнкт со значением *причинности*.

Следовательно, допустимым будет преобразование :

За зиму он по неосторожности сжег $[C_0]$ *машину дров* \Leftrightarrow *За зиму он по неосторожности уничтожил* $[Oper_1(S_{instr}(C_0))]$ *огнем* $[S_{instr}(C_0)]$ *машину дров.*
Но :

За зиму он сверхнормативно $\xleftarrow{5}$ *сжег* $[C_0]$ *машину дров.*

не допускает равнозначного преобразования в :

За зиму он сверхнормативно $\xleftarrow{5}$ *уничтожил* $[Oper_1(S_{instr}(C_0))]$ *огнем* $[S_{instr}(C_0)]$
машину дров,

поскольку *сверхнормативно* при сказуемом *уничтожил* не выражает адьюНКТА со значением *причинности*, а имеет Лексическое Значение “оценка”.

Свойства теории сорта Лексического Значения слова.

Свойство 1. Теория сорта опорного существительного Именной Группы является описанием свойств объектов, принадлежащих описываемому сорту. Практически данное свойство означает первостепенность отношений, непосредственно задаваемых опорными словами Именных Групп.

Свойство 2. Наличие операторов сортового сдвига дает возможность отдельно описывать λ -выражений для бинарных отношений, задаваемых Именными Группами.

Свойство 3. Структура утверждений-аксиом теории : отдельное утверждение задает бинарное отношения между понятиями известных Семантических Классов. Имя самого отношения задается словом с явно заданным для него Семантическим Классом.

Свойство 4. Задаваемое Именной Группой отношение рассматривается относительно сортов реалий, обозначаемых входящими в Именную Группу словами.

Информация теории Лексического Значения слова.

sort_theory (Sort, Var_Something_SemClass, Rel_list) (1)

Здесь списком *Rel_list* представляется набор утверждений (постулатов значений, meaning postulates) для данной теории.

Отдельный постулат-элемент списка *Rel_list* описывается структурой :

rel2 (Relation, Concept1, Concept2), (2)

где *Relation* – имя отношения, *Concept1* – управляющий, *Concept2* – зависимый член отношения.

Структура (2) является ссылкой на одно из известных Семантических Отношений. Сами Семантические Отношения представляются фактами динамической БД Visual Prolog'a :

relation2 (Relation, Concept1, Concept2), (3)

где аргументы *Concept1* и *Concept2* отношения *Relation* представляют собой Семантические Классы слов, обозначающих фигурирующие в отношении понятия, а само отношение задается словом с некоторым известным Семантическим Классом.

Информация теории Лексического Отношения.

relation_theory(SemClass, VarName_for_Concept1, (4)
VarName_for_Concept2, Rel_list)

В процессе интерпретации структуры (4) посредством набора утверждений из списка *Rel_list* устанавливается истинность отношения, задаваемого значением *SemClass*. К нему происходит сдвиг Семантического Класа *VarName_for_Concept1* опорного слова Именной Группы¹. Семантические Класы членов Именной Группы определяются значениями переменных *VarName_for_Concept1* и *VarName_for_Concept2*.

Конкретизация переменных, входящих в описания теорий, осуществляется посредством занесения в динамическую БД фактов :

concrete_definition(Var_Something, Var_Meaning), (5)

где *Var_Something* - фигурирующая в утверждениях теории переменная, *Var_Meaning* - ее значение. В процессе обработки списка *Rel_list* факты (5) используются для установления соответствия значений переменных в структурах (2) значениям соответствующих аргументов предиката (3) в известных фактах семантических отношений между понятиями.

¹ Это же значение становится в результате сортового сдвига значением Семантического Класа всего словосочетания Именной Группы.

Принадлежность слова заданному сорту и распознавание Лексических Отношений.

Правило для установления принадлежности слова Семантического Класа *Concept* заданному сорту *Sort* :

```
semantic_orientation_by_sort_theory (Concept, Sort) :-  
    d_sort_theory (sort_theory (Sort, Var_Something, Relation_Definition)),  
    assertz (concrete_definition (Var_Something, Concept)),  
    relation_processing (Relation_Definition),  
    retractall (concrete_definition(_, _)).
```

Правило для установления распознавания отношения *Rel*, задаваемого реляционной Именной Группой, на основе Семантических Класов опорного и зависимого слова :

```
can_be_in_relation2_for_noun_phrase (Rel, X, Y) :-  
    d_relation_theory (  
        relation_theory (Rel, Var_Something, Var_Substance, Relation_Definition)),  
        assertz (concrete_definition (Var_Something, X)),  
        assertz (concrete_definition (Var_Substance, Y)),  
        relation_processing (Relation_Definition), retractall (concrete_definition(_, _)).
```

Здесь утверждениями :

```
d_sort_theory (sort_theory (Sort, Var_Something_SemClass, Rel_list)) (6)
```

и

```
d_relation_theory (relation_theory (SemClass, VarName_for_Concept1,  
    VarName_for_Concept2, Rel_list)) (7)
```

динамической базы фактов Visual Prolog'a задаются описания теорий Лексических Значений и задаваемых ими отношений (Лексических Отношений).

Распознавание в тексте реляционной Именной Группы для заданного отношения.

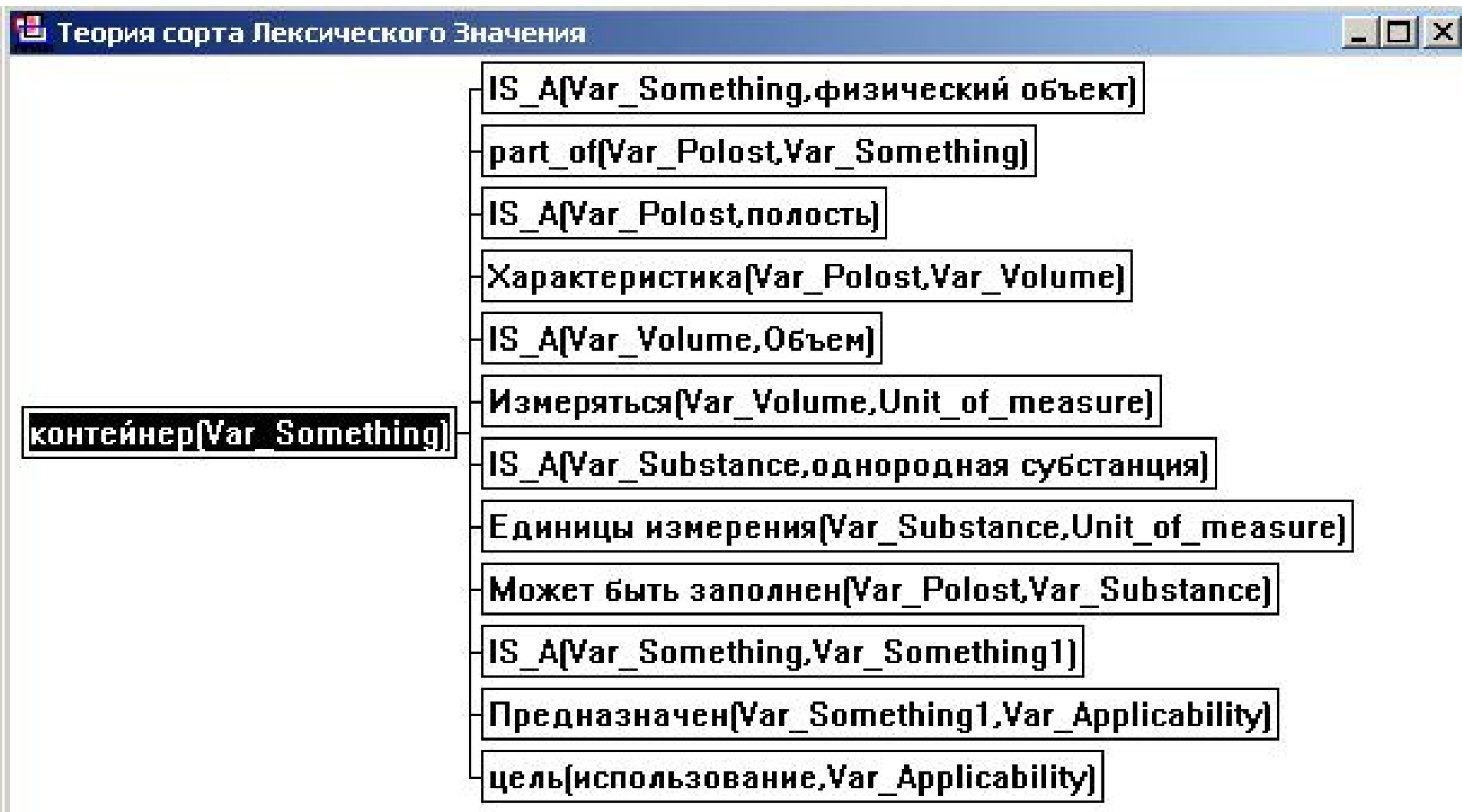
```
relational_noun_phrase (Rel, Word1, Word2) :-  
    morfol_characteristics (Word1, SyntClass1, GramInfo1),  
    morfol_characteristics (Word2, SyntClass2, GramInfo2),  
    has_a_relational_type (SyntClass1, GramInfo1, SyntClass2, GramInfo2),  
    semantic_orientation (Word1, SemClass1),  
    semantic_orientation (Word2, SemClass2),  
    can_be_in_relation2_for_noun_phrase (Rel, SemClass1, SemClass2).
```

Здесь *morfol_characteristics* – вспомогательное правило для определения Синтаксических Классов (*SyntClass1* и *SyntClass2*) и кодов грамматических характеристик (*GramInfo1* и *GramInfo2*) опорного *Word1* и зависимого *Word2* слова Именной Группы;

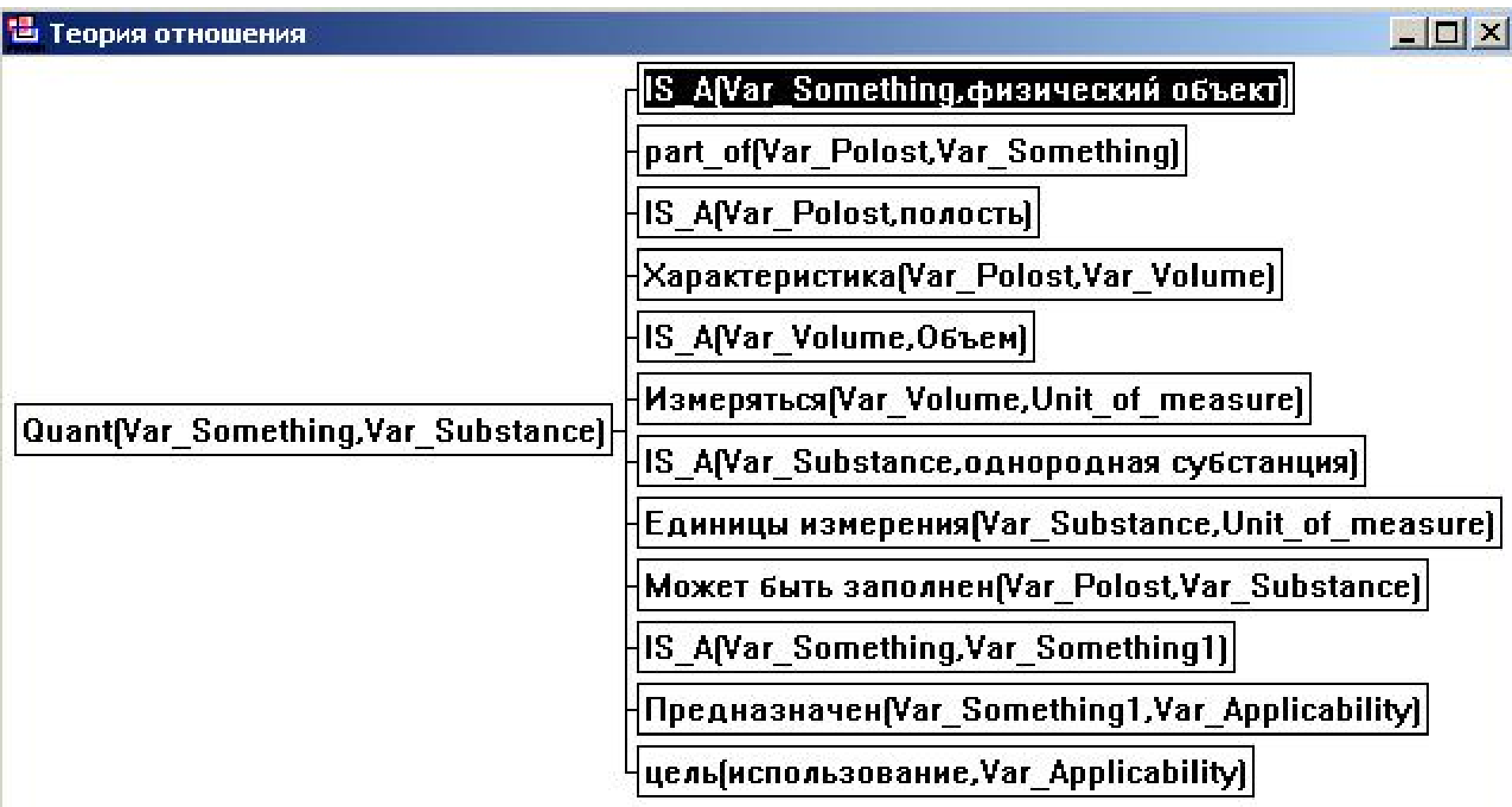
semantic_orientation (*Word*, *SemClass*) – вспомогательное правило для установления принадлежности слова *Word* одному из известных Семантических Классов *SemClass* с учетом отношения включения.

Посредством утверждений *has_a_relational_type* в динамическую базу фактов Visual Prolog'a заносится информация об известных видах реляционных Именных Групп.

Пример : теория Лексического Значения “контейнер”.



Теория отношения “квант”.



Формальный контекст для совокупности теорий Лексических Значений.

$$K=(G,M,W,I) \quad (8)$$

Здесь :

Множество объектов G есть множество слов-представителей заданного сорта (включая слово для обозначения самого сорта);

Множество формальных атрибутов M есть множество функций вида (9, 10), задаваемых утверждениями теорий и характеризующих обозначаемые словами понятия внутри данного сорта;

Множество атрибутивных значений W есть множество значений функций;

Тернарное отношение I между G , M и W , которое задает частичное отображение G на W : $m(g)=w$, ставит в соответствие каждой функции ее значение для рассматриваемого слова.

Примеры функций для вычисления характеристик ключевых свойств понятий.

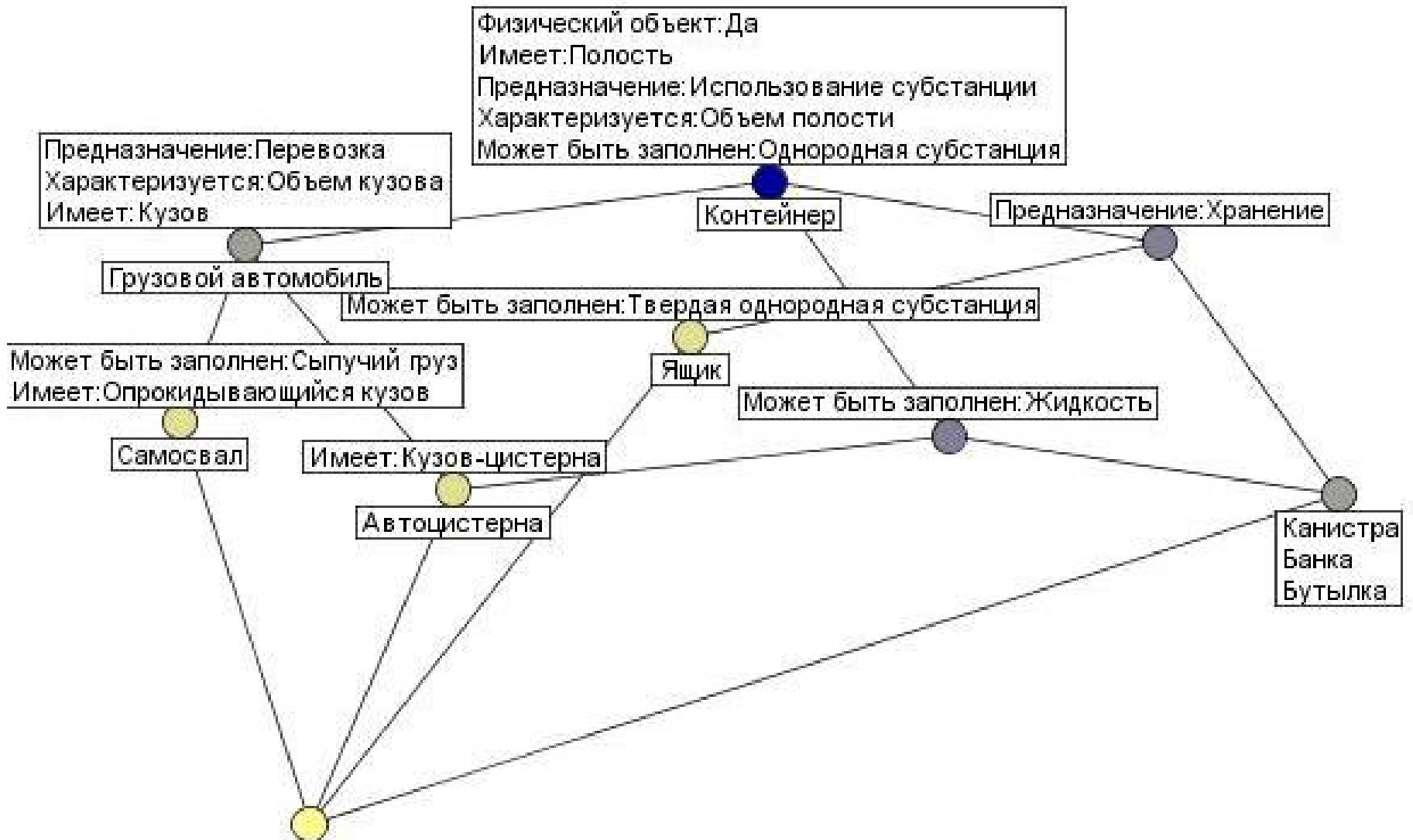
Функция для вычисления названия-этикетки характеристики “объем” для части объекта, которая может рассматриваться как “полость” :

$$\lambda z \lambda x [\exists y (\text{part_of}(x)(y) \ \& \ \text{полость}(y) \ \& \ \text{характеристика}(y)(z) \ \& \ \text{объем}(z))] \quad (9)$$

Применением принципа лямбда-конверсии Черча получаем выражение для абстрактного объекта “контейнер” :

$$\begin{aligned} & \lambda z \lambda x [\exists y (\text{part_of}(x)(y) \ \& \ \text{полость}(y) \ \& \ \text{характеристика}(y)(z) \ \& \ \text{объем}(z))] (\text{контейнер})(z) = \\ & \lambda z [\exists y (\text{part_of}(\text{контейнер})(y) \ \& \ \text{полость}(y) \ \& \ \text{характеристика}(y)(z) \ \& \ \text{объем}(z))] \end{aligned} \quad (10)$$

Модель системы теорий слов-представителей сорта "контейнер".



Выводы.

- **Разработанная методика представления и использования информации о Лексическом Значении слова в виде теории Семантического Класа (сорта) обозначаемого им объекта позволяет при построении дерева глубинного синтаксиса выбирать вариант установления отношений подчинения между словами в соответствии с требованиями Модели Управления предикатного слова для случая, когда Семантическим Класом актанта определяется некоторое отношение.**
- **Применительно к рассматриваемому случаю семантической ориентации глубинного синтаксического актанта предикатного слова введение в рассмотрение предложенного в настоящей работе формализованного описания теории сорта обозначаемого актантом объекта реальности позволяет оценить применимость значительного числа синонимических преобразований глубинной синтаксической структуры уже в процессе ее построения.**
- **Предложенная методика систематизации теорий-описаний Лексических Значений слов позволяет построить модель системы теорий в виде многозначного контекста применительно ко всему множеству сортов, для теорий которых характерно непустое пересечение множеств аксиом.**

Перспективы дальнейших исследований.

- Разработка модели лексического синонимического преобразования Глубинной Синтаксической Структуры с ориентацией на предложенную в настоящей работе методику представления и использования информации о лексическом значении слова.
- Применение модели лексического синонимического преобразования Глубинной Синтаксической Структуры при построении выражений *К-языков* для описания Семантического Представления текстов. При этом привлечение знаний о лексической синонимии позволяет ограничить объем информации Лексико-Семантического словаря описанием базовой лексики.
- Изучение возможностей построения формализованных теорий для контекстных отношений, не задаваемых напрямую Лексическим Значением опорного существительного Именной Группы.